V Namratha
1NT19IS180
C2 batch

```java
package nam;

import java.io.IOException;
import java.util.Iterator;
import java.util.StringTokenizer;

import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapred.FileInputFormat;
import org.apache.hadoop.mapred.FileOutputFormat;
import org.apache.hadoop.mapred.JobClient;
import org.apache.hadoop.mapred.JobConf;
import org.apache.hadoop.mapred.MapReduceBase;
import org.apache.hadoop.mapred.Mapper;
import org.apache.hadoop.mapred.OutputCollector;
import org.apache.hadoop.mapred.Reducer;
import org.apache.hadoop.mapred.Reporter;
import org.apache.hadoop.mapred.TextInputFormat;
import org.apache.hadoop.mapred.TextOutputFormat;

public class WordCount {

    public static class Map extends MapReduceBase implements
    Mapper<LongWritable, Text, Text, IntWritable> {
    private final static IntWritable one = new IntWritable(1);
    private Text word = new Text();
    public void map(LongWritable key, Text value, OutputCollector<Text, IntWritable>
    output, Reporter reporter) throws IOException {
    String line = value.toString();
    StringTokenizer tokenizer = new StringTokenizer(line);
    while (tokenizer.hasMoreTokens()) {
    word.set(tokenizer.nextToken());
    output.collect(word, one);
    }
```

```java
        }
    }

    public static class Reduce extends MapReduceBase implements
    Reducer<Text, IntWritable, Text, IntWritable> {
    public void reduce(Text key, Iterator<IntWritable> values,
    OutputCollector<Text, IntWritable> output, Reporter reporter)
    throws IOException {
    int sum = 0;
    while (values.hasNext()) {
    sum += values.next().get();
    }
    output.collect(key, new IntWritable(sum));
    }
    }

    public static void main(String[] args) throws Exception
    {
        JobConf conf = new JobConf(WordCount.class);
        conf.setJobName("wordcount");
        conf.setOutputKeyClass(Text.class);
        conf.setOutputValueClass(IntWritable.class);
        conf.setMapperClass(Map.class);
        conf.setCombinerClass(Reduce.class);
        conf.setReducerClass(Reduce.class);
        conf.setInputFormat(TextInputFormat.class);
        conf.setOutputFormat(TextOutputFormat.class);
        FileInputFormat.setInputPaths(conf, new Path(args[0]));
        FileOutputFormat.setOutputPath(conf, new Path(args[1]));
        JobClient.runJob(conf);



    }

}
```

```
hdoop@admin1-HP-280-G4-MT-Business-PC:~$ cd $HADOOP_HOME/sbin
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hdoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [admin1-HP-280-G4-MT-Business-PC]
Starting resourcemanager
Starting nodemanagers
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ jps
5474 SecondaryNameNode
5235 DataNode
2101 org.eclipse.equinox.launcher_1.5.600.v20191014-2022.jar
6165 Jps
5078 NameNode
5814 NodeManager
5662 ResourceManager
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -mkdir -p /input180
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -appendToFile - /input180 /text.txt
appendToFile: stdin (-) must be the sole input argument when present
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -appendToFile - /input180/text.txt
ghe who must not be named
hhe is voldemort2022-06-18 12:16:51,080 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted =
false
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hadoop jar /home/hdoop/Documents/namratha.jar /input180 /output180
2022-06-18 12:18:38,126 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2022-06-18 12:18:38,584 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2022-06-18 12:18:38,750 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute
 your application with ToolRunner to remedy this.
2022-06-18 12:18:39,030 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hdoop/.staging/job_16555346450
81_0001
2022-06-18 12:18:39,215 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-18 12:18:39,333 INFO mapred.FileInputFormat: Total input files to process : 1
2022-06-18 12:18:39,368 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-18 12:18:39,417 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
```

```
appendToFile: stdin (-) must be the sole input argument when present
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -appendToFile - /input180/text.txt
ghe who must not be named
hhe is voldemort2022-06-18 12:16:51,080 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted =
false
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hadoop jar /home/hdoop/Documents/namratha.jar /input180 /output180
2022-06-18 12:18:38,126 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2022-06-18 12:18:38,584 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2022-06-18 12:18:38,750 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute
 your application with ToolRunner to remedy this.
2022-06-18 12:18:39,030 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hdoop/.staging/job_16555346450
81_0001
2022-06-18 12:18:39,215 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-18 12:18:39,333 INFO mapred.FileInputFormat: Total input files to process : 1
2022-06-18 12:18:39,368 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-18 12:18:39,417 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-18 12:18:39,434 INFO mapreduce.JobSubmitter: number of splits:2
2022-06-18 12:18:39,526 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2022-06-18 12:18:39,584 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1655534645081_0001
2022-06-18 12:18:39,584 INFO mapreduce.JobSubmitter: Executing with tokens: []
2022-06-18 12:18:39,882 INFO conf.Configuration: resource-types.xml not found
2022-06-18 12:18:39,882 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2022-06-18 12:18:40,084 INFO impl.YarnClientImpl: Submitted application application_1655534645081_0001
2022-06-18 12:18:40,113 INFO mapreduce.Job: The url to track the job: http://admin1-HP-280-G4-MT-Business-PC:8088/proxy/application_1655534645081_0001
/
2022-06-18 12:18:40,114 INFO mapreduce.Job: Running job: job_1655534645081_0001
2022-06-18 12:18:45,229 INFO mapreduce.Job: Job job_1655534645081_0001 running in uber mode : false
2022-06-18 12:18:45,229 INFO mapreduce.Job:  map 0% reduce 0%
2022-06-18 12:18:49,300 INFO mapreduce.Job:  map 100% reduce 0%
2022-06-18 12:18:56,439 INFO mapreduce.Job:  map 100% reduce 100%
2022-06-18 12:18:58,470 INFO mapreduce.Job: Job job_1655534645081_0001 completed successfully
2022-06-18 12:18:58,608 INFO mapreduce.Job: Counters: 54
        File System Counters
                FILE: Number of bytes read=104
                FILE: Number of bytes written=677556
                FILE: Number of read operations=0
```

```
        File System Counters
                FILE: Number of bytes read=104
                FILE: Number of bytes written=677556
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=247
                HDFS: Number of bytes written=62
                HDFS: Number of read operations=11
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
                HDFS: Number of bytes read erasure-coded=0
        Job Counters
                Launched map tasks=2
                Launched reduce tasks=1
                Data-local map tasks=2
                Total time spent by all maps in occupied slots (ms)=3685
                Total time spent by all reduces in occupied slots (ms)=4117
                Total time spent by all map tasks (ms)=3685
                Total time spent by all reduce tasks (ms)=4117
                Total vcore-milliseconds taken by all map tasks=3685
                Total vcore-milliseconds taken by all reduce tasks=4117
                Total megabyte-milliseconds taken by all map tasks=3773440
                Total megabyte-milliseconds taken by all reduce tasks=4215808
        Map-Reduce Framework
                Map input records=2
                Map output records=9
                Map output bytes=80
                Map output materialized bytes=110
                Input split bytes=182
                Combine input records=9
                Combine output records=9
                Reduce input groups=9
                Reduce shuffle bytes=110
                Reduce input records=9
                Reduce output records=9
```

```
                Total vcore-milliseconds taken by all map tasks=3685
                Total vcore-milliseconds taken by all reduce tasks=4117
                Total megabyte-milliseconds taken by all map tasks=3773440
                Total megabyte-milliseconds taken by all reduce tasks=4215808
        Map-Reduce Framework
                Map input records=2
                Map output records=9
                Map output bytes=80
                Map output materialized bytes=110
                Input split bytes=182
                Combine input records=9
                Combine output records=9
                Reduce input groups=9
                Reduce shuffle bytes=110
                Reduce input records=9
                Reduce output records=9
                Spilled Records=18
                Shuffled Maps =2
                Failed Shuffles=0
                Merged Map outputs=2
                GC time elapsed (ms)=160
                CPU time spent (ms)=1510
                Physical memory (bytes) snapshot=774516736
                Virtual memory (bytes) snapshot=7611191296
                Total committed heap usage (bytes)=828375040
                Peak Map Physical memory (bytes)=292495360
                Peak Map Virtual memory (bytes)=2534359040
                Peak Reduce Physical memory (bytes)=191168512
                Peak Reduce Virtual memory (bytes)=2544308224
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
```

hdoop@admin1-HP-280-G4-MT-Business-PC: ~/hadoop-3.2.1/sbin

```
                    Shuffled Maps =2
                    Failed Shuffles=0
                    Merged Map outputs=2
                    GC time elapsed (ms)=160
                    CPU time spent (ms)=1510
                    Physical memory (bytes) snapshot=774516736
                    Virtual memory (bytes) snapshot=7611191296
                    Total committed heap usage (bytes)=828375040
                    Peak Map Physical memory (bytes)=292495360
                    Peak Map Virtual memory (bytes)=2534359040
                    Peak Reduce Physical memory (bytes)=191168512
                    Peak Reduce Virtual memory (bytes)=2544308224
            Shuffle Errors
                    BAD_ID=0
                    CONNECTION=0
                    IO_ERROR=0
                    WRONG_LENGTH=0
                    WRONG_MAP=0
                    WRONG_REDUCE=0
            File Input Format Counters
                    Bytes Read=65
            File Output Format Counters
                    Bytes Written=62
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ hdfs dfs -cat /output180/part*
2022-06-18 12:19:20,552 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
be          1
ghe         1
hhe         1
is          1
must        1
named       1
not         1
voldermort          1
who         1
hdoop@admin1-HP-280-G4-MT-Business-PC:~/hadoop-3.2.1/sbin$ []
```